



SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: Xu, Wenfeng
Presnell, Scott R.
Yee, David P.
Foster, Donald C.
- (ii) TITLE OF THE INVENTION: PROTEASE-ACTIVATED RECEPTOR
PAR4 (ZCHEMR2)
- (iii) NUMBER OF SEQUENCES: 12
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: ZymoGenetics, Inc.
 - (B) STREET: 1201 Eastlake Avenue East
 - (C) CITY: Seattle
 - (D) STATE: WA
 - (E) COUNTRY: USA
 - (F) ZIP: 98102
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette
 - (B) COMPUTER: IBM Compatible
 - (C) OPERATING SYSTEM: DOS
 - (D) SOFTWARE: FastSEQ for Windows Version 2.0
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Leith, Debra K
 - (B) REGISTRATION NUMBER: 32,619
 - (C) REFERENCE/DOCKET NUMBER: 98-10

*Sub
CS*

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 206-442-6674

(B) TELEFAX: 206-442-6678

(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4895 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence

(B) LOCATION: 176...1330

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTCCACGGG CTGGCTGGCA AGCGGCCCTG GTGGGTCTGC GGGGGCAGGG GCAGCCTTCC	60
TGGTTTATCT CCACCGGCGC GATCTGCTCG TCCGCCTCGG CTCCAGAAGC TGGGGCTCAG	120
GGTCCGGCGA GGCAGGAAGC CTGAGGCCAC AGCCCAGAGC AGCCTGAGTG CAGTC ATG	178
Met	
1	
TGG GGG CGA CTG CTC CTG TGG CCC CTG GTG CTG GGG TTC AGC CTG TCT	226
Trp Gly Arg Leu Leu Trp Pro Leu Val Leu Gly Phe Ser Leu Ser	
5 10 15	
GGC GGC ACC CAG ACC CCC AGC GTC TAC GAC GAG AGC GGG AGC ACC GGA	274
Gly Gly Thr Gln Thr Pro Ser Val Tyr Asp Glu Ser Gly Ser Thr Gly	
20 25 30	
GGT GGT GAT GAC AGC ACG CCC TCA ATC CTG CCT GCC CCC CGC GGC TAC	322
Gly Gly Asp Asp Ser Thr Pro Ser Ile Leu Pro Ala Pro Arg Gly Tyr	
35 40 45	
CCA GGC CAA GTC TGT GCC AAT GAC AGT GAC ACC CTG GAG CTC CCG GAC	370
Pro Gly Gln Val Cys Ala Asn Asp Ser Asp Thr Leu Glu Leu Pro Asp	
50 55 60 65	

AGC	TCA	CGG	GCA	CTG	CTT	CTG	GGC	TGG	GTG	CCC	ACC	AGG	CTG	GTG	CCC	418
Ser	Ser	Arg	Ala	Leu	Leu	Leu	Gly	Trp	Val	Pro	Thr	Arg	Leu	Val	Pro	
			70						75						80	
GCC	CTC	TAT	GGG	CTG	GTC	CTG	GTG	GTG	GGG	CTG	CCG	GCC	AAT	GGG	CTG	466
Ala	Leu	Tyr	Gly	Leu	Val	Leu	Val	Val	Gly	Leu	Pro	Ala	Asn	Gly	Leu	
			85						90					95		
GCG	CTG	TGG	GTG	CTG	GCC	ACG	CAG	GCA	CCT	CGG	CTG	CCC	TCC	ACC	ATG	514
Ala	Leu	Trp	Val	Leu	Ala	Thr	Gln	Ala	Pro	Arg	Leu	Pro	Ser	Thr	Met	
			100				105					110				
CTG	CTG	ATG	AAC	CTC	GCG	ACT	GCT	GAC	CTC	CTG	CTG	GCC	CTG	GCG	CTG	562
Leu	Leu	Met	Asn	Leu	Ala	Thr	Ala	Asp	Leu	Leu	Leu	Ala	Leu	Ala	Leu	
			115				120					125				
CCC	CCG	CGG	ATC	GCC	TAC	CAC	CTG	CGT	GGC	CAG	CGC	TGG	CCC	TTC	GGG	610
Pro	Pro	Arg	Ile	Ala	Tyr	His	Leu	Arg	Gly	Gln	Arg	Trp	Pro	Phe	Gly	
130						135				140					145	
GAG	GCC	GCC	TGC	CGC	CTG	GCC	ACG	GCC	GCA	CTC	TAT	GGT	CAC	ATG	TAT	658
Glu	Ala	Ala	Cys	Arg	Leu	Ala	Thr	Ala	Ala	Leu	Tyr	Gly	His	Met	Tyr	
				150					155					160		
GGC	TCA	GTG	CTG	CTG	CTG	GCC	GCC	GTC	AGC	CTG	GAT	CGC	TAC	CTG	GCC	706
Gly	Ser	Val	Leu	Leu	Leu	Ala	Ala	Val	Ser	Leu	Asp	Arg	Tyr	Leu	Ala	
			165					170					175			
CTG	GTG	CAC	CCG	CTG	CGG	GCC	CGC	GCC	CTG	CGT	GGC	CGG	CGC	CTG	GCC	754
Leu	Val	His	Pro	Leu	Arg	Ala	Arg	Ala	Leu	Arg	Gly	Arg	Arg	Leu	Ala	
			180				185					190				
CTT	GGA	CTC	TGC	ATG	GCT	GCT	TGG	CTC	ATG	GCG	GCC	GCC	CTG	GCA	CTG	802
Leu	Gly	Leu	Cys	Met	Ala	Ala	Trp	Leu	Met	Ala	Ala	Ala	Leu	Ala	Leu	
			195				200					205				
CCC	CTG	ACA	CTG	CAG	CGG	CAG	ACC	TTC	CGG	CTG	GCG	CGC	TCC	GAT	CGC	850
Pro	Leu	Thr	Leu	Gln	Arg	Gln	Thr	Phe	Arg	Leu	Ala	Arg	Ser	Asp	Arg	
210						215				220				225		
GTG	CTC	TGC	CAT	GAC	GCG	CTG	CCC	CTG	GAC	GCA	CAG	GCC	TCC	CAC	TGG	898
Val	Leu	Cys	His	Asp	Ala	Leu	Pro	Leu	Asp	Ala	Gln	Ala	Ser	His	Trp	
				230					235					240		

CAA CCG GCC TTC ACC TGC CTG GCG CTG TTG GGC TGT TTC CTG CCC CTG	946
Gln Pro Ala Phe Thr Cys Leu Ala Leu Leu Gly Cys Phe Leu Pro Leu	
245 250 255	
CTG GCC ATG CTG CTG TGC TAC GGG GCC ACC CTG CAC ACG CTG GCG GCC	994
Leu Ala Met Leu Leu Cys Tyr Gly Ala Thr Leu His Thr Leu Ala Ala	
260 265 270	
AGC GGC CGG CGC TAC GGC CAC GCG CTG AGG CTG ACC GCA GTG GTG CTG	1042
Ser Gly Arg Arg Tyr Gly His Ala Leu Arg Leu Thr Ala Val Val Leu	
275 280 285	
GCC TCC GCC GTG GCC TTC TTC GTG CCC AGC AAC CTG CTG CTG CTG CTG	1090
Ala Ser Ala Val Ala Phe Phe Val Pro Ser Asn Leu Leu Leu Leu Leu	
290 295 300 305	
CAT TAC TCG GAC CCG AGC CCC AGC GCC TGG GGC AAC CTC TAT GGT GCC	1138
His Tyr Ser Asp Pro Ser Pro Ser Ala Trp Gly Asn Leu Tyr Gly Ala	
310 315 320	
TAC GTG CCC AGC CTG GCG CTG AGC ACC CTC AAC AGC TGC GTG GAT CCC	1186
Tyr Val Pro Ser Leu Ala Leu Ser Thr Leu Asn Ser Cys Val Asp Pro	
325 330 335	
TTC ATC TAC TAC TAC GTG TCG GCC GAG TTC AGG GAC AAG GTG CGG GCA	1234
Phe Ile Tyr Tyr Tyr Val Ser Ala Glu Phe Arg Asp Lys Val Arg Ala	
340 345 350	
GGG CTC TTC CAA CGG TCG CCG GGG GAC ACC GTG GCC TCC AAG GCC TCT	1282
Gly Leu Phe Gln Arg Ser Pro Gly Asp Thr Val Ala Ser Lys Ala Ser	
355 360 365	
GCG GAA GGG GGC AGC CGG GGC ATG GGC ACC CAC TCC TCT TTG CTC CAG T	1331
Ala Glu Gly Gly Ser Arg Gly Met Gly Thr His Ser Ser Leu Leu Gln	
370 375 380 385	
GACACAAAGT GGGGAAGGCT GTACTGGGTC GAACAGGGTC CCTTCCCCCA CTTACGTCC	1391
TTCCTGGGAC CTCAGAATGT GACCTTATTT GGAAATAGGG TTGTTACAAC TGTCAGTAGC	1451
GGAGGTCACT TTGGAGAAGG GTGGGCCTTA CATCCAGTGT GGGTGGTGTC CTCATAAGAT	1511
AAGGAGAGGC CAGGCCTGGT GGCTCACGCC TGTAATCCCA GCACTTTAAG AGGCCAAGGC	1571
GGATGGATCA CTTGAGCCCA GGAGTTCAAC ACCAGCCTGA GCAACATGGT AAAACCCCAT	1631
CTCTACCAAA AATACAAAAA TTAGCTGGGC TTGGTGGCTG GCGCCTGTAA TCCAGCTAC	1691
TCAGGAGACT GAGGCAGAAG GATCGCTTGA ACCTGGGAGG CAGAGGTTGC AGTGAGCCGA	1751
GATTGCGCCA CTGGACTCCA GCCTGCGTGA CAGAGAGCCT GTCTCTAAAT TAATTAATTA	1811
ATTAATTTAA TTCAATTTTA AAAAGACGAA AAGTGACGGC CAGGTGCAGT GGCTCACGCC	1871

TATAATCTCA	GCACTCTGGG	AGGCCAAGAT	GGAGGATTGC	TTGAAGCCAG	GAGTTTGGGA	1931
CCAGCCTGGG	CAACATAGGG	GGATCCCATC	TCTACACACA	AAAAAATTTT	TTAATGAACC	1991
AGGCATTGTG	GCATGCGCCT	ATAGTCCCAG	CCACTCAAGA	GGCACAGGCG	GGAGGATCAC	2051
TTGAGCCTGG	GAGGTTGTGG	TTGCAGTGAG	CTATGATTGT	ACCACTGCAC	TCCAGCCTGG	2111
GCAACAGAGC	AAGACCTTGT	CTCAAAAATA	AACAACTAA	AATTAAAAAA	AGAAGACGAG	2171
AGATAGTGGG	TGTGGTGGCT	CACACCTGCA	ATCCCAGCAC	TTTGAAGGC	CGAGGTGGGC	2231
AGATCATCTG	AGGCCAGGAG	TTCAAGACCA	GCCTGGCTAA	CATGGTGAAA	TCCTATCTCT	2291
ACCAAAAATA	CAAAAATTAG	CCAGGCGTGG	TGGTGGGCAC	CTGTACTGGG	GAGGTGCCCA	2351
CCCAGCTACT	GGGGAGGCTG	AGTCAGGAGA	ATCGCTTGAA	CCTGGGAGGC	GGAGGTTGCG	2411
GTCAGCTGAG	ATGGTGCCAC	TGCACTCCAG	CCTGGGCGAA	AGAGCGACTC	TGTCTCCAAA	2471
AAAAAGAGAA	GAGGAGAGGA	CACAGAGACA	CACAGAGAAG	AAAGCCATGT	GGCGGCAGAG	2531
GCAGAGATGG	GAGTGATGCG	GACGGACACA	AACTAAGGGA	TGCCACGATG	CCAAGCACAG	2591
CCAACAGCCA	CCAGCAGCCA	GGAGACAGGC	CTGGGACGGG	CTCTCCCTCA	CAGCCTCCAG	2651
AGGGAACCG	CCCTGCCACC	ACCTTGACCC	TGGACTTCTG	GCCTGCAGAA	CTGTGAGACA	2711
ATAAACTCTC	ATTGTTTTAA	GCTGCCTGGC	ATGTGGCACT	TTGTCAGGGC	AGCCCAGGAA	2771
TCTGAAACAG	GATCAAACCT	TGCTTCCTGG	GCCCTGCCAG	CATCTCTGGC	TCGGCTTTCT	2831
GGGCTGGATG	CAGCCCACGA	CGCACTGGTG	TCTGAGATGG	GGCTGGAGCT	GGGGCTGGGG	2891
CTGCATTCCC	TGGAGACTCA	CTGCAAGTTC	CTGCCCAGGA	GGCTGAGGGC	ACCCCATCCT	2951
CAGTGCCCAA	TGCTGTGGCC	CCACCAGGCC	CAGAGCCTGG	TTGGCCATTC	TCATGCCCAC	3011
CAGCTTCTGG	CTTTGGGATG	TCTCTTGAGC	AACCAGAATA	GCACCCCAA	CTCTGCTCCC	3071
CAAAACCCAT	CACTAGCACG	GCTCAGCCTC	CTGCTATCCC	CTGACTGCTG	GGGACCCTCG	3131
CCTTCCCTCC	TCTCACCTGC	AGGCTGATCC	TTCTTTTCAC	TTTCTGTCAA	TGTCACCAGG	3191
GATAAGGTGG	GACAATGGGG	GGTGGGGGTG	GACAGTGTGT	GCTGGGGGGT	TCGGGTGCTG	3251
CAGACCTGGA	ACTCCCTTCT	GCCAGGATGT	TGGCAGCCGG	TTGTAAGCCT	TGCACGGGAC	3311
AGACCACACC	CACCGCAACC	TCATCCCCCT	AGCACTAACC	ACATCCACTC	TCAACCCCGT	3371
CCCCTTCGCA	CTGACCACAC	CCACCCCGTT	CGGCCCCGCC	CCCCGCACTG	AACACTCCCG	3431
CCCTCAACCC	CGCACCCCTCC	GCACTCACCT	CCCCCTCGCC	GCTCGACCCC	GCCCTCACCA	3491
CACTGACCAC	CCTCAACCCA	TTGCGCCCAG	TCCCCACCAC	AGTGACCACA	CCCTCACTGG	3551
CTCGGCCCTG	CCCCCAGTAT	ACTGACCATT	CCCCAGCCAC	TTCCCTTCCG	CACCTACCAC	3611
TCCCCAGCC	ACGCCCCCTCC	CCGCTGACCG	CTCCTCCAGC	CCCGCCTCCC	CCGTACAGGC	3671
AGAGCGCCCG	CCCACCTCTA	TGCTGCGTTC	TCCTGACTTT	ACGTTGGCCC	CTCCTCTGCC	3731
AAGCCCCCAG	GGGAGCCCTC	CCTGGCGTCC	GAGGGTGGGA	GTCGGGGTGT	GGCAGGCCGC	3791
GGTGGGGGGC	GGCAGTGGCT	CCGCGCACTC	ACCCGGGGCC	CGGGCAGGGG	CGCGCTCCAC	3851
TTCGTTGCAC	GCGGGTCCGG	CGCACAGTTC	CCGGGCGAGT	GGGCTGTGCG	TGCTGACGTT	3911
GTAGAAGCGA	GTGGCCTCGA	AGGCTACGGG	ACGAGGGTGG	CGGGTGACCA	AGTGCAGGCG	3971
CGACGGGTCA	GGGACCGGGC	CGGGCCGGGG	GTGCGGGCGC	GCGGGCCTAC	CGGGTTCGTA	4031
GTAGTCGTAC	ACGGAGACTG	GCAGCGCCGA	CGTCCTGCC	ACCACGCACT	CCCGGAGAGC	4091
ACGGAACCGC	ACGCACGTCA	GGCACCGGCT	GGGGATCTGT	GGGGCAGCGG	CGGGCGCAGG	4151
CTCGACCCGG	GCCAGGAGGC	CCGGGGCGCT	GAGCTCAGGC	CCAGAACTGG	CTGATTTTCT	4211
GGATACCCAG	GACGCGTGAA	ACACAGAAGA	AACGTGATCC	CATTTTCTTT	TTTTCTTTTA	4271
CTTTTCTTTT	TTTTTTTTTT	TCCTGAGACA	GAGTCTCGCG	CTGTTGCCCA	GGCTGGAGTG	4331
CAGTGGCGTG	ATCTCGGCTC	ACTGCAAGCT	CGGCCTCCTG	GGTTCAAATG	ATTCTCCTGC	4391
CTCAGCCTCC	CAAGTAGCTG	GGATAACAGG	CGCCACCAAC	CGCACCTGCT	TAATTTTTTG	4451
TATTTTTGAT	CAAGACGGAG	TTTACCATG	TTGGCCAGGC	TGGTCTCCAA	CTCCTGCCCT	4511
CAAGTGATCC	GCCTCGGTCC	CATTTTTTAT	TCTTTGGGTC	CTTCCATCCC	ACTGGGAAAA	4571

```

CGTCTCAGGT GGCCTCTGAA ACACCACTCC TTTTGTGTG TGTGCACGCA TGGCTGAGCA 4631
TGTGTGGGTG GGAGTCAGCA CATTACGAT ACTGTGCAAT CATCACCTCT GTCTAGTTAC 4691
AGGACGGTTT CTTTCTCCCC CAAAGAAACC CCATCGCCAT CAGCACTCAC TCCCCACTCC 4751
CCCAGCCCCT GGCAACCACA AATCTTTCCA ACTCTACGGA TTTGCCTGTT CTGGGCATTT 4811
CATGTCAATG GAATCATGTA CTCTGTGAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 4871
AAAAAAAAA AAAAAAAAAA AAAA 4895

```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 385 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Trp Gly Arg Leu Leu Leu Trp Pro Leu Val Leu Gly Phe Ser Leu
 1           5           10           15
Ser Gly Gly Thr Gln Thr Pro Ser Val Tyr Asp Glu Ser Gly Ser Thr
 20           25           30
Gly Gly Gly Asp Asp Ser Thr Pro Ser Ile Leu Pro Ala Pro Arg Gly
 35           40           45
Tyr Pro Gly Gln Val Cys Ala Asn Asp Ser Asp Thr Leu Glu Leu Pro
 50           55           60
Asp Ser Ser Arg Ala Leu Leu Leu Gly Trp Val Pro Thr Arg Leu Val
 65           70           75           80
Pro Ala Leu Tyr Gly Leu Val Leu Val Val Gly Leu Pro Ala Asn Gly
 85           90           95
Leu Ala Leu Trp Val Leu Ala Thr Gln Ala Pro Arg Leu Pro Ser Thr
 100          105          110
Met Leu Leu Met Asn Leu Ala Thr Ala Asp Leu Leu Leu Ala Leu Ala
 115          120          125
Leu Pro Pro Arg Ile Ala Tyr His Leu Arg Gly Gln Arg Trp Pro Phe
 130          135          140
Gly Glu Ala Ala Cys Arg Leu Ala Thr Ala Ala Leu Tyr Gly His Met
 145          150          155          160
Tyr Gly Ser Val Leu Leu Leu Ala Ala Val Ser Leu Asp Arg Tyr Leu
 165          170          175
Ala Leu Val His Pro Leu Arg Ala Arg Ala Leu Arg Gly Arg Arg Leu
 180          185          190

```

Ala Leu Gly Leu Cys Met Ala Ala Trp Leu Met Ala Ala Ala Leu Ala
 195 200 205
 Leu Pro Leu Thr Leu Gln Arg Gln Thr Phe Arg Leu Ala Arg Ser Asp
 210 215 220
 Arg Val Leu Cys His Asp Ala Leu Pro Leu Asp Ala Gln Ala Ser His
 225 230 235 240
 Trp Gln Pro Ala Phe Thr Cys Leu Ala Leu Leu Gly Cys Phe Leu Pro
 245 250 255
 Leu Leu Ala Met Leu Leu Cys Tyr Gly Ala Thr Leu His Thr Leu Ala
 260 265 270
 Ala Ser Gly Arg Arg Tyr Gly His Ala Leu Arg Leu Thr Ala Val Val
 275 280 285
 Leu Ala Ser Ala Val Ala Phe Phe Val Pro Ser Asn Leu Leu Leu Leu
 290 295 300
 Leu His Tyr Ser Asp Pro Ser Pro Ser Ala Trp Gly Asn Leu Tyr Gly
 305 310 315 320
 Ala Tyr Val Pro Ser Leu Ala Leu Ser Thr Leu Asn Ser Cys Val Asp
 325 330 335
 Pro Phe Ile Tyr Tyr Tyr Val Ser Ala Glu Phe Arg Asp Lys Val Arg
 340 345 350
 Ala Gly Leu Phe Gln Arg Ser Pro Gly Asp Thr Val Ala Ser Lys Ala
 355 360 365
 Ser Ala Glu Gly Gly Ser Arg Gly Met Gly Thr His Ser Ser Leu Leu
 370 375 380
 Gln
 385

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1155 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGTGGGGNM GNYTNYTNYT NTGGCCNYTN GTNYTNGGNT TYWSNYTNWS NGGNGGNACN	60
CARACNCCNW SNGTNTAYGA YGARWSNGGN WSNACNGGNG GNGGNGAYGA YWSNACCCN	120
WSNATHYTNC CNGCNCNMG NGGNTAYCCN GGNCARGTNT GYGCNAAYGA YWSNGAYACN	180
YTNGARYTNC CNGAYWSNWS NMGNGCNYTN YTNYTNGGNT GGGTNCNAC NMGNYTNGTN	240
CCNGCNYTNT AYGGNYTNGT NYTNGTNGTN GGNYTNCNG CNAAYGGNYT NGCNYTNTGG	300
GTNYTNGCNA CNCARGCNC NMGNYTNCCN WSNACNATGY TNYTNATGAA YYTNGCNACN	360
GCNGAYYTNY TNYTNGCNYT NGCNYTNCCN CCMGNATHG CNTAYCAYYT NMGNGGNCAR	420

MGNTGGCCNT	TYGGNGARGC	NGCNTGYMGN	YTNGCNACNG	CNGCNYTNTA	YGGNCAYATG	480
TAYGGNWSNG	TNYTNYTNYT	NGCNGCNGTN	WSNYTNGAYM	GNTAYYTNGC	NYTNGTNCAY	540
CCNYTNMGNG	CNMGNGCNYT	NMGNGGNMGN	MGNYTNGCNY	TNGGNYTNTG	YATGGCNGCN	600
TGGYT NATGG	CNGCNGCNYT	NGCNYTNCCN	YTNACNYTNC	ARMGN CARAC	NTTYMGNYTN	660
GCNMGNWSNG	AYMGNGTNYT	NTGYCAYGAY	GCNYTNCCNY	TNGAYGCNCA	RGCNWSNCAY	720
TGGCARCCNG	CNTTYACNTG	YYTNGCNYTN	YTNGGNTGYT	TYYTNCNNYT	NYTNGC NATG	780
YTNYTNTGYT	AYGGNGCNAC	NYTNCAYACN	YTNGCNGCNW	SNGGNMGNMG	NTAYGGNCAY	840
GCNYTNMGNY	TNACNGCNGT	NGTNYTNGCN	WSNGCNGTNG	CNTTYTTYGT	NCCNWSNAAY	900
YTNYTNYTNY	TNYTNCAyta	YWSNGAYCCN	WSNCCNWSNG	CNTGGGGNAA	YYTNTAYGGN	960
GCNTAYGTNC	CNWSNYTNGC	NYTNWSNACN	YTNAAYWSNT	GYGTNGAYCC	NTTYATHTAY	1020
TAYTAYGTNW	SNGCNGARTT	YMGNGAYAAR	GTNMGNGCNG	GNYTNTTYCA	RMGNWSNCCN	1080
GGNGAYACNG	TNGCNWSNAA	RGCNWSNGCN	GARGGNNGNW	SNMGNGGNAT	GGGNACNCAY	1140
WSNWSNYTNY	TNCAR					1155

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Ile Thr Thr Cys His Asp Val
1 5

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Asp Ser Lys Gly Ser Ser Gln Lys Gly Ser Arg Leu Leu Leu Leu
1 5 10 15
Leu Val Val Ser Asn Leu Leu Leu Cys Gln Gly Val Val Ser Asp Tyr
20 25 30

Lys Asp Asp Asp Asp Lys Leu Glu Gly Gly
 35 40

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Asp Tyr Lys Asp Asp Asp Asp Lys
 1 5

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Gly Tyr Pro Gly Gln Val
 1 5

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Ser Phe Leu Leu Arg Asn

1

5

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TGGCACTGCC CCTGACACTG CA

22

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CCCGTAGCAC AGCAGCATGG

20

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGTGCCCGCC CTCTATGG

18

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TCGCGAGGTT CATCAGCA